DOCUMENT RESUME

ED 334 236 TM 016 761

AUTHOR Curlette, William L.; And Others
TITLE An Evaluation Model To Select an Integrated Learning System in a Large, Suburban School District.
PUB DATE Apr 91
NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 3-7, 1991).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; *Computer Assisted Instruction; Computer Software Evaluation; Computer Software Selection; Courseware; Decision Making; Elementary Education; Evaluation Criteria; *Evaluation Methods; *Models; *School Districts; Suburban Schools
IDENTIFIERS *DeKalb School District GA; *Integrated Learning Systems; Multi Attribute Utility Theory

ABSTRACT

The systematic evaluation process used in Georgia's DeKalb County School System to purchase comprehensive instructional software--an integrated learning system (ILS)--is described, and the decision-making model for selection is presented. Selection and implementation of an ILS were part of an instructional technology plan for the DeKalb schools that called for common hardware and computers in the classrooms. The ILS was to be introduced to grades 2 through 6 over a period of 5 years. Initially, a committee of instructional leaders and local school personnel rated 13 software packages from several vendors. Three systems--Computer Curriculum Corporation, Education System Corporation, and World Institute of Computer Assisted Teaching (WICAT)--were tested in three elementary schools using third grade students (N=229). The Multi-Attribute Utility Method from decision theory was used to evaluate these systems without regard to cost. The evaluation took into account: (1) student achievement while using the ILS; (2) teacher experience in using the ILS; and (3) technical and administrative experience with the software. The analysis indicated that all three systems more than met the school system's criteria. At that point, cost was considered, leading to the selection of the system of the WICAT. The decision-making model went beyond preevaluation to test the software in classroom situations, yielding vital information for the school system. One table is included. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED334236

AN EVALUATION MODEL TO SELECT AN INTEGRATED LEARNING SYSTEM IN A LARGE, SUBURBAN SCHOOL DISTRICT

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

William L. Curlette
Georgia State University
Educational Foundations

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
William Curlette
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Katrena C. Howard and Su Ellen W. Bray
DeKalb County School System
Research & Evaluation Department
Dunwoody, Georgia

Paper presented to the American Education Research Association
on
April 5, 1991
Chicago, Illinois

TM016761

BEST COPY AVAILABLE

AN EVALUATION MODEL TO SELECT AN INTEGRATED LEARNING SYSTEM IN A LARGE, SUBURBAN SCHOOL DISTRICT

The decision-making process used by a school system to purchase comprehensive instructional software is typically limited to reading descriptions of the product and observing demonstrations of the software. Although this pre-evaluation is necessary, it stops short of providing sufficient information on which to base decisions which involve millions of dollars.

In Georgia's DeKalb County School System, however, personnel followed a systematic evaluation process for selecting an ILS that went beyond pre-evaluation and included a) student achievement when using different software products, b) actual teacher experience in using the software with students, and c) technical and administrative experiences with the software. Another unique aspect of DeKalb's evaluation process was separating the quality issue from the cost until the last step.

Although much has been published on the positive impact of computer instruction, only a few studies compare vendors with one another and only one study (Resta and Rost, 1986) was located that compared the results for decision-making. Numerous articles, however, describe criteria to consider when selecting an Integrated Learning System and/or a process to follow (Wilson, 1990; Education Turnkey).

Resta and Rost reported their preliminary results of a comparative study of four ILS's used with Chapter I students. The study was designed to focus on important instructional, technical, and performance characteristics of the ILS packages, the conditions for their effective use, and the correlation of system educational content with district curriculum. Resta and Rost looked at 17 categories including match with the district curriculum and standardized tests used in the district, instructional characteristics, program uses, management and record keeping and

reporting, technical characteristics, training, cost, achievement, and student/teacher perceptions. Although the Resta and Rost study is relevant to the study described in this report, a major difference was that the computers were placed in a lab whereas the computers in this study were placed in each classroom. Another study conducted by Mitzel (1970) also stressed the importance of including both pupil achievement and pupil feelings as part of comprehensive evaluation of Computer Assisted Instruction (CAI).

Researchers in New York (Swan, Mitrani, Gurerrero, Schoener, 1989) conducted a large scale study of 13 computer based instructional programs in 26 elementary and secondary schools. The study examined the efficiency of using comprehensive computer-based instruction for providing basic skills remediation to educationally disadvantaged student populations. Thus, the study did not compare performance among vendors. Three of the 13 programs were the same programs piloted in the school system in this study.

Overall achievement results from the New York study indicated that computer instruction was an effective means for delivering instruction, was as effective in providing instruction in reading as in math to the educationally disadvantaged, and was more effective for the younger students than for the older students. When comparing reading scores of third grade students using the three programs that were piloted in DeKalb, students using WICAT experienced the most gains followed by students using ESC and then CCC. In mathematics achievement, students using CCC had the most gains followed by students using WICAT and then ESC.

A major objective of this paper is to present a decision-making model for selecting instructional software which included a new statistical approach for summarizing quantitative data. This paper addresses specific issues concerning the effects of technology on student learning, the evaluation of school-based programs, and the improvement of data gathering in

school settings. The Integrated Learning Systems (ILS) evaluated in this study represent the state of the art in computer assisted instruction currently offered by three of the nation's largest vendors. In addition to describing how personnel reached the decision to purchase one vendor's software, results concerning each of the three major software products in the study are reported.

Background Information

The DeKalb County School System is located east of Atlanta and has a student population of over 73,000, the largest in the state. The selection and implementation of an ILS was part of DeKalb's five year Instructional Technology Plan. Under this plan, four important selection process decisions had been made by the school system prior to selecting ILS's to participate in the evaluation process.

The first decision was that all ILS software would be run on common hardware in the pilot so that differences could not be attributed to equipment. The second decision was that four terminals would be placed in each classroom (distributive mode) instead of a lab. There were several reasons for placing the computers within the classroom. Primarily, educational computing personnel believe that if technology is to be integrated successfully into the teaching process, then technology tools must be physically present in the classroom. Also, the use of the distributive mode addressed the issues of efficient use of teacher and student time.

The third decision was that an ILS for reading and math instruction would be implemented in all 230 third grade classrooms of 76 elementary schools, and the implementation would be expanded by at least one grade a year and, therefore, would be present in all second through sixth grade classrooms at the end of a five year period dependent upon an annual review and Board of Education approval. The fourth decision was that the ILS would individualize

instruction to meet the needs of all students: remedial, average, and gifted. Furthermore, the ILS would be used to enrich student learning as well as to develop skills.

Methods and Procedures

Educational Computing personnel, along with evaluators in the Research & Evaluation Department, worked with a research methodologist in designing the program evaluation and in collecting and analyzing the data. In deciding to implement an ILS and in conducting the evaluation, evaluators gave attention to the following goals:

1. to increase student achievement,
2. to assist teachers in individualizing instruction, and
3. to improve student skills in thinking critically and solving problems.

The first step in implementation was to form a committee composed of instructional leaders and local school personnel who reviewed available ILS's. The committee, working with consultants from the Department of Educational Computing, conducted an investigation of ILS's by reviewing the literature, and by contacting and/or visiting school systems throughout the country that use ILS's. The committee then rated software from several vendors using the following criteria:

- content match with the school system's curriculum,
- quality,
- instructional impact,
- support material, and
- management system.

Based upon this review and evaluation of the instructional software and the management system in relation to the school system's needs, the committee identified the following leading contenders which met the specifications at a sufficiently high level to warrant being in the pilot project: Computer Curriculum Corporation (CCC), Education System Corporation (ESC, later to become Josten's), and World Institute of Computer Assisted Teaching (WICAT).

A different ILS was placed in the third grades of three elementary schools. Although all third grade students (N=229) in each school used an ILS for the entire school year, the pilot study assessed the overall robustness and instructional effectiveness of the ILS using an eight-week period. Each vendor installed equipment in a pilot school before the school year began, and vendor representatives trained teachers to implement the system.

Deciding upon Categories and Weights

Edwards, Guttentag, and Snapper (1975) derived the Multi-Attribute Utility (MAUT) Method from decision-theory which has been used for program planning. The basic purpose of MAUT is to separate the elements of a complicated decision and evaluate each element separately. Rather than approaching a decision globally and beginning a debate about the merits of competing programs, decision-makers begin by reaching agreement on what they are trying to accomplish through available programs. Having agreed on overall goals, they then develop a consensus on the criteria to be used in judging the desirability of the competing programs. Decision-makers then weight the relative importance of the criteria, evaluate each program on the basis of each criterion, and combine the evaluation to form an overall judgment to show the utility or effectiveness of each program. By discussing the values and agreeing on priorities at the outset, decision-makers can determine whether a conflict of values exist or whether a conflict exists on how well each program was judged to fulfill a common set of values.

Using the MAUT Method, the evaluators and program implementors selected ten categories to use in evaluating the ILS's. An expert panel comprised of teachers, principals, and instructional coordinators rated and reached a consensus on the categories to be used in judging the desirability of the competing programs. The panel then assigned a weight to each category basing decisions not only on the relative importance of each but also on the quantity and quality of data that could be obtained to score each category. For example, because only a short questionnaire assessed student perceptions, a weight of four was assigned. However, even though panel members were aware that student achievement data would be based on only eight weeks of instruction, they still decided that this category should have a relatively high weight in case differences were found. The resulting eight categories and weights are listed below:

	CATEGORY	WEIGHT
1.	Curriculum Match	10
2.	Achievement	9
3.	Ease of Use by Teachers	8
4.	Track Record of Vendor	7
5.	Quality of Training	6
6.	Teachers', Administrators', Coordinators' Opinions of Effectiveness	5
7.	Students' Attitudes about How Much Computer Helped Them to Learn	4
8.	Corporate Stability of Vendor	3

Purposely, the evaluation process did not include cost as one of the weighted categories. While data from the evaluation were being analyzed, Department of Educational Computing personnel wrote specifications for a cost proposal. The department's two-year experience with the pre-evaluation process prior to the pilot study and their experience during the study with teacher training, testing various cabling networking configurations, and working with the management programs in the schools enabled the department to address this school system's needs in the cost specifications. When the bids were opened, the cost of each ILS system was compared to the evaluation scores.

Data Sources

Data were collected from numerous sources such as records and documents (e.g. student printouts), student, teacher and administrator questionnaires, observations, standardized achievement tests, and system-developed tests. Table 1 details the data sources used for each of the eight categories.

In addition, teachers and system attendants in each school documented the amount of time each student used the ILS. This documentation, along with classroom observations and meetings with the teachers using each ILS, verified that the ILS was implemented consistently in each school and classroom. A number of stakeholders who provided input into the evaluation included instructional coordinators, teachers, administrators, and central office personnel. All of the data were triangulated where possible.

Combining Scores

A step in the evaluation model was to arrive at an overall score to summarize the quantitative data obtained during the data collection. This summary was accomplished using the two T-scores presented by Kiresuk and Sherman (1968) in which raw scores are converted to

TABLE 1

DATA SOURCES FOR EACH WEIGHTED CATEGORY

CATEGORY	SOURCES
1. Curriculum Match	Reading Coordinator Math Coordinator Math Coordinator-Instructional Considerations Teachers-Reading and Math
2. Achievement	Reading on Iowa Math on Iowa plus additional items
3. Ease of Use	System Attendants Teachers
4. Track Record	Department of Educational Computing System Attendants Teachers
5. Training	Third Grade Teachers and System Attendants: Standard Training Questionnaires 1st and 2nd System Attendants Teachers
6. Instructional Effectiveness	Teachers Administrators
7. Student Attitudes	Students
8. Corporate Stability	Department of Educational Computing

z-scores. The original T-score approach assumes a common intercorrelation of a particular value between the indicators. A new methodological feature which was an outgrowth of this research is to iterate the correlation from 0 to 1 by increments of .10 to see if the overall rank ordering of the three vendors change. No change in the rank within a reasonable range of possible correlation provides evidence that the procedure is robust regarding Kiresuk's and Sherman's recommended correlation of .30. To some extent, the use of this formula under these circumstances expands its application since it was designed primarily for use in clinical settings.

Results

The statistical aspects of the evaluation model provided overall numerical scores for each vendor that at least allowed rank ordering. In this case, the analysis of the data indicated that all three vendors more than met the school system's criteria for an ILS, and the cost of the product became the significant factor. The recommendation to the Board of Education for a five million dollar ILS to be implemented over the next five years was for the vendor who ranked second on the T-score evaluation because their overall cost was for approximately three million dollars less than the other two competitors. Thus, the recommendation was for WICAT. As the decision-maker, the Board of Education approved the recommendation.

Discussion

The pilot of an ILS in actual classrooms yielded vital information to school system personnel. This direct experience provided information about each vendor's quality of service from their local field representatives. System personnel identified strengths of each vendor, and those strengths were used in writing specifications for a cost proposal and in easing the first year's system-wide implementation, particularly in the area of training teachers. In addition, teachers were able to experiment with various classroom management strategies which became

invaluable to other teachers when the program was implemented system-wide the following year. The pilot study gave curriculum leaders time to confirm whether or not and to what extent the vendor's curriculum supported the school system's. By using an ILS for a year, teachers had time to accept this new instructional tool. Because of the additional information provided by the pilot study, the year of experience with the pilot proved to be even more valuable than originally anticipated.

This decision-making model goes beyond pre-evaluation to test the effectiveness of different software programs in actual school settings. This model can provide personnel in other school systems with a way to decide upon a program, particularly an expensive Integrated Learning System. Most importantly, this evaluation model, which separated cost from quality, saved the school system over a million dollars in the first year's implementation alone. Furthermore, the addition of the new iteration procedure to the modified T-score provides an improved method for obtaining overall scores.

References

Education Turnkey Systems, Inc. (1989, May). Study of integrated learning systems/instructional networks. Presented at the Software Publishers Association Annual Spring Conference, San Diego, CA.

Edwards, Guttentag, & Snapper (1975). A decision-theoretic approach to evaluation research. In Struening and Guttentag (Eds.), Handbook of Evaluation Research (vol. 1). Newbury Park: Sage.

Kiresuk, T.J. & Sherman, R.E. (1968). Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. Community Mental Health Journal, 4, 443-453.

Mitzel, H.E. (1970). How to evaluate computer assisted instruction. University Park, PA: Penn State University. (ERIC Document Reproduction Service No. ED 047 005)

Resta, P.E. & Rost, P. (1986, April). CAI: A model for the comparison and selection of integrated learning systems in large school districts. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Swan, K., Mitrani, M., Guerrero, F., & Schoener, J. (1989, June). The use of comprehensive computer-based instruction for the basic skills remediation of educationally disadvantaged students. Paper presented at the National Educational Computing Conference, Boston, MA.

Wilson, J. (1990). Integrated learning systems: A primer. Classroom Computer Learning, 10, 22-23, 27-30, 34,36.